

# Unit 24: Confidence Intervals



## SUMMARY OF VIDEO

This video is an introduction to inference, which means we use information from a sample to infer something about a population. For example, we might use a sample statistic to estimate a population parameter. Suppose we wanted to know a man's mean blood pressure. A sample of blood pressure readings is shown in Table 24.1.

Su	M	T	W	Th	F	Sa	
130	120	140	125	130	130	140	A.M.
125	130	145	140	125	135	110	P.M.

*Table 24.1. Systolic blood pressure readings.*

We could estimate his mean blood pressure using the sample mean from these readings,  $\bar{x} = 130$ . But how trustworthy is our conclusion given that different samples could lead to different results, some higher and others lower? Statisticians address this issue by calculating confidence intervals. Rather than a single number like 130, we can compute a range of values along with a confidence level for that range.

Next, the context switches from blood pressure to the length of life of batteries. Because companies promise specific battery lifetimes and improved performance over a competitor, they need proof before ads promoting their product go on the air. At Kodak's Ultra Technologies, technicians use rigorous testing and calculate confidence intervals to back up their marketing claims. Here's how the data are collected. Random samples of batteries are pulled from the warehouse. The batteries are drained under controlled conditions and the time it takes for them to run out of juice is recorded. From these data, Kodak has determined that its population of AA batteries when used in a toy will last  $7\frac{1}{2}$  hours  $\pm$  20 minutes and that their confidence in that range is 95%.

Now, we retrace Kodak's steps to figure out how they came up with this interval. Before getting started, we need to check that a few underlying assumptions are satisfied:

1. Observations are independent.
2. Data are from a normal population or the sample size  $n$  is large.
3. The population standard deviation is known.

Selecting a random sample of batteries for the test takes care of the assumption of independent observations. The second assumption is satisfied since the sample size of  $n = 40$  is considered large. The last assumption is not reasonable in the real world, but for now, we'll assume that from past data we do know the population standard deviation,  $\sigma = 63.5$  minutes.

The task is to calculate a confidence interval for  $\mu$ , the mean life of Kodak's AA batteries. Our sample statistic  $\bar{x}$  is a point estimate for the parameter  $\mu$ . If we include a margin of error around our point estimate, we get an interval estimate of the form:

point estimate  $\pm$  margin of error

From Unit 22, Sampling Distributions, we know that the sampling distribution of  $\bar{x}$  is normal, with mean  $\mu_{\bar{x}} = \mu$  and standard deviation  $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ . In this case, we are given  $\sigma = 63.5$  minutes and we can compute  $\sigma_{\bar{x}} = 63.5/\sqrt{40}$  minutes or about 10 minutes. Think back to the 68-95-99.7% Rule. In any normal distribution, 95% of the observations lie within two standard deviations of the mean. So, 95% of all possible samples result in battery-life data for which  $\mu$  is within plus or minus 20 minutes of that sample's mean,  $\bar{x}$ . In our example,  $\bar{x} = 450$  minutes. So, we can say with 95% confidence that  $\mu$  lies within 20 minutes of  $\bar{x}$ , giving us a confidence interval from 430 minutes to 470 minutes. To say that we are 95% confident in our calculated range of values means that we got the numbers using a method that gives correct results 95% of the time over many, many examples.

What if Kodak were willing to settle for only 90% confidence? Or what if they insisted on 99% confidence? We can get any confidence level that we want by turning to the standard normal distribution and finding the  $z^*$  critical value. Then just substitute the appropriate values into the following formula:

$$\bar{x} \pm z^* \left( \frac{\sigma}{\sqrt{n}} \right)$$

Notice that the margin of error gets larger if we insist on higher confidence because  $z^*$  will be larger. On the other hand, the margin of error gets smaller if we take more observations so that  $n$  is larger.

# STUDENT LEARNING OBJECTIVES

- A. Understand that a common reason for taking a sample is to estimate some property of the underlying population.
- B. Recognize that a useful estimate requires a measure of how accurate the estimate is.
- C. Know that a confidence interval has two parts: an interval that gives the estimate and the margin of error, and a confidence level that gives the likelihood that the method will produce correct results in the long range.
- D. Be able to assess whether the underlying assumptions for confidence intervals are reasonably satisfied. Provided the underlying assumptions are satisfied, be able to calculate a confidence interval for  $\mu$  given the sample mean, sample size, and population standard deviation.
- E. Understand the tradeoff between confidence and margin of error in intervals based on the same data.
- F. Given a specific confidence level, recognize that increasing the size of the sample can give a margin of error as small as desired.

# CONTENT OVERVIEW

The purpose of a confidence interval is to estimate an unknown parameter with an indication of (1) how precise the estimate is and (2) how confident we are that the result is correct. Any confidence interval has two parts: an interval computed from the data and a confidence level. The interval often has the form

point estimate  $\pm$  margin of error.

The **confidence level** states the probability that the method will give a correct result. That is, if you use 95% confidence intervals often, in the long run 95% of your intervals will contain the true parameter value.

Suppose that a simple random sample of size  $n$  is drawn from a normally distributed population having an unknown mean  $\mu$  and known standard deviation  $\sigma$ . A level  $C$  (expressed as a decimal) confidence interval for  $\mu$  is

$$\bar{x} \pm z^* \left( \frac{\sigma}{\sqrt{n}} \right),$$

where  $z^*$  is a cutoff point for the standard normal curve with area  $(1 - C)/2$  to its right. For example, if  $C = 0.95$  (for a 95% confidence interval) then  $(1 - C)/2 = (1 - 0.95)/2 = 0.025$ . In this case,  $z^*$  turns out to be 1.96 as shown in Figure 24.1.

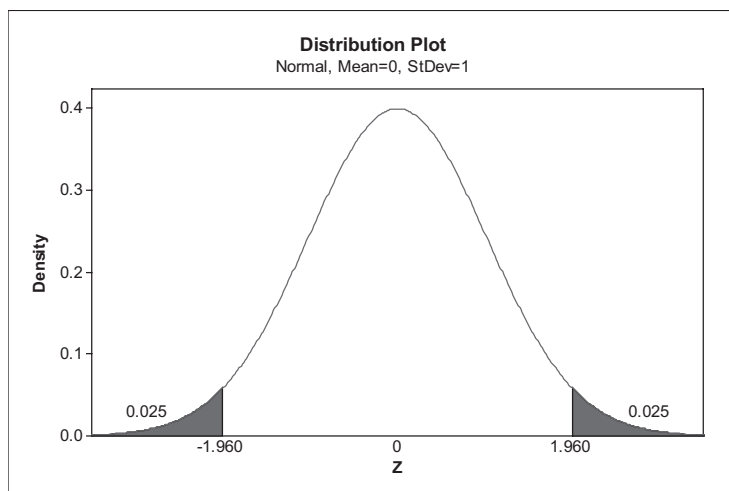


Figure 24.1. Standard normal density curve illustrating  $z^* = 1.96$ .

If the sample size  $n$  is relatively small, we first need to check that the underlying assumption of normality is reasonably satisfied before computing a confidence interval. One way to check the assumption of normality is to make a normal quantile plot of the sample data. Alternatively,

we could look at a boxplot. If the sample size  $n$  is large ( $n$  at least 30), the confidence interval formula is approximately correct even when the population does not have a normal distribution. This result is due to the Central Limit Theorem (Unit 22, Sampling Distributions).

The size of the margin of error controls the precision (width) of the confidence interval estimate. Precision is increased as the margin of error shrinks. The margin of error of a confidence interval decreases if any of the following occur:

- the confidence level  $C$  is reduced
- the sample size  $n$  increases
- the population standard deviation  $\sigma$  decreases

In practice, the population standard deviation  $\sigma$  is not known and must be estimated from the sample. If the sample size  $n$  is fairly large (say at least 30), then the value of the sample standard deviation  $s$  should be close to  $\sigma$ . In that case, you can replace  $\sigma$  by  $s$  in the confidence interval formula. (See Unit 26, Small Sample Inference for One Mean, for a continued discussion of confidence intervals for  $\mu$  when  $\sigma$  is unknown.)

# KEY TERMS

A **point estimate** of an unknown population parameter is a single number based on sample data (a statistic) that represents a plausible value for that parameter.

A **confidence interval** for a population parameter is an interval of plausible values for that parameter. It is constructed so that the value of the parameter will be captured between the endpoints of the interval with a chosen level of confidence. The **confidence level** is the success rate of the method used to construct the confidence interval.

Many confidence intervals have the following form: point estimate  $\pm$  margin of error. The **margin of error** is the range of values above and below the point estimate.

A formula used to compute a **confidence interval for  $\mu$**  when  $\sigma$  is known and either the sample size  $n$  is large or the population distribution is normal is given by:

$$\bar{x} \pm z^* \left( \frac{\sigma}{\sqrt{n}} \right),$$

where  $z^*$  is a z-critical value associated with the confidence level.

# THE VIDEO

Take out a piece of paper and be ready to write down answers to these questions as you watch the video.

1. Why is a single blood pressure reading not sufficient if we want to estimate a person's average blood pressure?
2. What are the two parts of any confidence interval?
3. What assumptions need to be checked before computing a confidence interval?
4. In plain language, what does "95% confidence" mean?

# UNIT ACTIVITY:

## CONFIDENCE INTERVAL SIMULATION

In this activity, you will need the simulation data collected for question 2 in Unit 22's activity. Recall that samples of size 9 were drawn from an approximately normal distribution with  $\mu = 50$  and standard deviation  $\sigma = 4$ . Assume for the moment that  $\mu$  is unknown. You will be using sample data to find confidence interval estimates for  $\mu$ .

1.
  - a. What is the standard deviation of  $\bar{x}$  for samples of size 9?
  - b. What is the margin of error for a 95% confidence interval for  $\mu$ ? (Round your answer to two decimals.)
2. Your instructor has a container filled with numbered strips. Draw a sample of size 9.
  - a. Record the outcomes of your sample and calculate the sample mean,  $\bar{x}$ .
  - b. Based on your sample, determine a 95% confidence interval for  $\mu$ .
  - c. In this case, the true value of  $\mu$  is 50. Does your confidence interval contain the true value of  $\mu$ ?
3. Your instructor should distribute a table of the results from 100 samples of size 9 generated for Unit 22's activity.
  - a. For each sample, calculate a 95% confidence interval estimate for  $\mu$  and record the endpoints of the interval.
  - b. Of the 100 samples collected, how many of the 95% confidence intervals contain the true value of  $\mu$ , which is 50? How many did you expect to contain 50? Is there a discrepancy between the number you found and the number you expected to find? Explain how this discrepancy could occur.



# EXERCISES

1. Students who take SATs in most high schools are not representative of all students in the school. Generally, only students who plan to apply for college admission take the test. The statistics class at Lincoln High decides to get better data. They select a random sample of 20 members of the junior class and arrange for all 20 to take the Math SAT.

The scores are given below.

410	400	460	440	390	400	450	460	520	380
480	480	490	450	480	330	390	460	600	610

Assume that the standard deviation of scores for all juniors is  $\sigma = 100$ .

- Find the value of  $\sigma_{\bar{x}}$ , the standard deviation of the sample mean in size-20 samples.
- Check to see whether these data could be considered to come from a normally distributed population. (The data need only be roughly normal – in other words, the data should have no severe departures from normality.)
- Let  $\mu$  be the mean score that would be observed if every junior at Lincoln High took the exam. Give a 95% confidence interval for  $\mu$ . Show your calculations. How could you get a smaller margin of error with the same confidence?
- Give a 99% confidence interval for  $\mu$ . Explain in plain language, to someone who knows no statistics, why this interval is wider than your result in (c).

2. The Massachusetts Comprehensive Assessment System (MCAS) includes a 10th grade math test which is scaled from 200 to 280. Assume that the standard deviation of math test scores is  $\sigma = 17$ . (This assumption is reasonable based on past results.)

- A random sample of 30 test results is given below. Use these results to determine a 95% confidence interval for the mean MCAS math score,  $\mu$ .

252	266	264	244	262	268	236	254	264	276
266	220	218	260	258	232	268	218	262	242
238	262	250	264	276	234	232	266	276	248

b. A second random sample of 30 test results was taken. The results are given below. Combine the data from the two samples, the one below and the one in (a), and use the combined data to compute a 95% confidence interval for  $\mu$ .

258	252	268	264	264	264	222	258	220	254
254	274	266	264	268	248	238	248	258	254
254	258	208	268	268	272	274	254	272	270

c. Compare the margin of errors for the confidence intervals in (a) and (b). Why would you expect the margin of error based on 60 observations to be less than the margin of error based on 30 observations?

d. Keeping the confidence level at 95%, how many observations would you need in order to reduce the margin of error to under 3.0?

3. A city planner randomly selects 100 apartments in Boston, Massachusetts, to estimate the mean living area per apartment. The sample yielded  $\bar{x} = 875$  square feet with a standard deviation  $s = 255$  square feet.

a. Calculate a 95% confidence interval for  $\mu$ , the mean living area per apartment. (Keep in mind that since the sample size is large,  $s$  should be close to  $\sigma$ .)

b. Having found the interval in (a), can you say there is a 95% chance that the mean living area is within the interval? Explain why or why not.

4. A random sample of 50 full-time, hourly wage workers between the ages of 20 and 40 was selected from participants in the 2012 March Supplement, which is part of the Current Population Survey (a joint venture of the U.S. Bureau of Labor Statistics and Census Bureau). The hourly rate (in dollars) of these workers is given below.

7.25	30.09	12.00	25.00	8.00	27.53	14.20	31.00	20.00	18.00
12.00	28.12	16.50	8.00	9.00	15.00	15.10	18.00	17.43	14.00
15.25	34.50	8.00	14.80	7.80	11.00	33.07	10.55	19.00	19.50
12.25	18.00	24.00	27.50	15.00	6.75	30.00	10.30	27.00	14.50
8.00	14.00	10.00	11.75	15.00	28.00	7.50	28.50	16.25	11.75

- a. Calculate the sample mean and standard deviation.
- b. Calculate a 95% confidence interval for  $\mu$ , the mean hourly wage of full-time, hourly wage workers between the ages 20 and 40. Because the sample size is large, use  $s$ , the sample standard deviation, in place of  $\sigma$ , the unknown population standard deviation.
- c. A politician speaking around the time that the data for the 2012 March Supplement were collected claimed that salaries were rising. He stated that the average hourly rate for fulltime workers between the ages of 20 and 40 was \$20.00. Does your confidence interval from (b) affirm or refute the politician's claim. Explain.
- d. After being confronted, the politician complained that we should have used a 99% confidence interval to estimate the mean hourly wage. Compute a 99% confidence interval for  $\mu$ . Does the 99% confidence interval affirm his claim? Explain.

# REVIEW QUESTIONS

1. There are many thousands of male high school basketball players. Julie collects the heights of the 96 varsity players in her school's league. The mean height of these 96 players is  $\bar{x} = 71.1$  inches and the standard deviation is  $s = 2.7$  inches.

a. Because the sample is large, the population standard deviation  $\sigma$  will be close to 2.7, the observed sample standard deviation. Give a 95% confidence interval for the mean height of all varsity basketball players, assuming that Julie's observations are a random sample. Show your calculations.

b. Do you think it is reasonable to take these 96 players as a random sample of all male varsity basketball players? Why or why not?

2. A random sample of 36 skeletal remains from females was taken from data stored in the Forensic Anthropology Data Bank (FDB) at the University of Tennessee. The femur lengths (right leg) in millimeters are recorded below.

432	432	435	460	432	440	448	449	434	443
525	451	448	443	450	467	436	423	475	435
433	438	453	438	435	413	439	442	507	424
468	419	434	483	448	514				

a. Determine the sample mean and standard deviation.

Since the sample size is large, we can use the sample standard  $s$  in place of  $\sigma$  in calculations of confidence intervals.

b. Before doing any calculations, think about a 90%, 95% and 99% confidence for  $\mu$ , the mean femur bone length for women. Which of these intervals would be the widest? Which would be the narrowest? Explain how you know without calculating the confidence intervals.

c. Calculate 90%, 95%, and 99% confidence intervals for  $\mu$ , the mean femur bone length for adult females. Do your results confirm your answer to (b)?

3. Birth weights in grams of a random sample of 20 babies born in Massachusetts in 2010 appear below. Based on past data, assume that the standard deviation for birth weights is  $\sigma = 600$  grams .

4054 3572 2636 3430 3118 3969 3628 3940 4536 4819

3883 3487 3827 3883 2749 3487 3855 4450 4309 3345

- a. Are the underlying assumptions for confidence intervals reasonably satisfied?
- b. Determine a 95% confidence interval for the mean birth weights of babies born in Massachusetts.
- c. In the United States, we are more accustomed to reporting babies' weights in ounces (or even pounds and ounces) than grams. How would you modify the confidence interval to give a confidence interval for the mean weight in ounces? Calculate that interval. (Use the following conversion: 1 gram  $\approx$  0.03527 ounce.) Does your result seem reasonable?

4. How much can a single outlier affect a confidence interval? Suppose that the first observation of 4054 grams in the random sample in question 3 had been 350 grams (the weight of a baby that did not survive).

- a. Make a boxplot of the modified data set to show that this low weight baby is an outlier.
- b. Recalculate the 95% confidence interval based on the modified data. How much did the outlier affect the confidence interval?

Final comment: Always look at your data before calculating confidence intervals. Outliers can greatly affect your results.